

Personalized Government Online Services with Recommendation Techniques

A Thesis

Submitted to the University Graduate School of
University of Technology, Sydney

by

Xuetao Guo

In Full Fulfillment of the Requirements for the
Degree of Doctor of Philosophy in Computing Science

2006

Certificate of Authorship/Originality

I certify that the work in this thesis has not previously been submitted for a degree and nor has it been submitted as part of requirements for a degree except as fully acknowledged within the text.

I also certify that the thesis has been written by me. Any help that received in my research work and the preparation of the thesis has been acknowledged. In addition, I certify that all information sources and literatures used are indicated in the thesis.

Signature of Candidate

Xuetao Guo

Date

08/06/2006

Acknowledgements

I would like to express my sincere thanks to all those who have provided me with various supports during my Ph.D study. My special thanks first go to my supervisor, Associate Professor Jie Lu. I am deeply indebted to her for her continuous support and encouragement over the last three years. Next, I would like to thank Associate Professor Simeon Simoff and Dr. Guangquan Zhang for their valuable advices and many useful discussions, which have greatly improved the completion of my Ph.D study. I would also like to acknowledgement Dr. Chris Cornelie for our interesting discussions about my study.

Finally, I am also very grateful to have a wonderful family. I thank their enduring love and wholehearted support throughout this study. Without saying, my sincerest thanks to my wife, Jennifer, and daughter, Ellen, for always being supportive, and for providing the necessary occasional distraction from my work, during the intense time this thesis was put together.

While working on my Ph.D study, I was supported by the Australian Postgraduate Award (APA) scholarship. Also, I thank the financial support from the Faculty of Information Technology, University of Technology Sydney during the study. The work contained in this thesis was carried out in the E-service Research Group, Faculty of Information Technology at the University of Technology Sydney.

Table of Contents

Certificate of Authorship/Originality	i
Acknowledgements	ii
Table of Figures	vi
List of Tables	viii
Abstract	ix
Chapter 1 Introduction	1
1.1 Background	2
1.2 Research problems	4
1.3 Research objectives	5
1.4 Significance	6
1.5 Contributions	7
1.6 Thesis structure	8
1.7 Summary	10
Chapter 2 Literature Review	11
2.1 Information retrieval and information filtering	12
2.1.1 Information overload	12
2.1.2 Information retrieval	12
2.1.3 Information filtering	15
2.2 Web personalization	17
2.2.1 Concept of web personalization	17
2.2.2 Taxonomy of web personalization	19
2.2.3 Design of personalization systems	30
2.3 Recommender systems	31
2.3.1 Concept of recommender systems	31
2.3.2 Existing recommendation approaches	32
2.3.3 Existing recommendation algorithms	48
2.3.4 Relevant techniques for recommendations	53
2.4 Fuzzy logic techniques for recommendation	63
2.5 Summary	66
Chapter 3 Research Methodology	68
3.1 Research methods and process	69
3.2 Relevant research methods used in this study	71

3.2.1	Qualitative and quantitative research methods.....	71
3.2.2	Evaluation research approach.....	72
3.2.3	Experimental and observational research.....	80
3.2.4	Prototyping approach of information systems.....	80
3.2.5	Appropriate approaches used in this study.....	81
3.3	Summary	82
Chapter 4 Development Model of Australian Government Online Services.....		84
4.1	Effectiveness evaluation of government online services.....	85
4.1.1	Design of an evaluation model	85
4.1.2	Data collection	90
4.1.3	Data analysis.....	91
4.2	Development model of government online services	97
4.2.1	Existing development models of government online services	98
4.2.2	Personalized model of government online services.....	99
4.3	Summary	101
Chapter 5 Hybrid Recommendation Approaches for One-and-only Items.....		103
5.1	One-and-only item recommendation	104
5.2	A hybrid recommendation approach integrated with semantic similarity analysis	105
5.2.1	Similarity and prediction computation.....	105
5.2.2	Hybrid recommendation algorithms with semantic similarity analysis	112
5.2.3	A numerical example of recommendation generation	117
5.3	A hybrid recommenadtion algorithm with fuzzy logic.....	119
5.4	Summary	125
Chapter 6 Smart Trade Exhibition Finder: System Design and Implementation		126
6.1	STEF: smart trade exhibition recommender system.....	127
6.2	Product taxonomy.....	128
6.3	Semantic product relevance model.....	130
6.4	Analysis of the STEF system	132
6.4.1	Conceptual framework of dynamic personalization	132
6.4.2	Data flow diagram of the STEF system	135
6.4.3	Overall procedure of recommendation in the system	136
6.5	STEF: Prototype architecture.....	140

6.6	STEF: System implementation.....	145
6.7	Summary	147
Chapter 7 Experimental Evaluation and Result Analysis		148
7.1	Experiment design	149
7.1.1	Data sets and experimental procedure	149
7.1.2	Evaluation metrics	150
7.2	Experimental results	152
7.2.1	Optimal neighbourhood size.....	152
7.2.2	Optimal β value	153
7.2.3	Comparison of recommendation accuracy	154
7.2.4	Impacts of different recommendation approaches.....	155
7.2.5	Discussion of the results.....	157
7.3	Summary	158
Chapter 8 Conclusions and Further Study.....		160
8.1	Conclusions	161
8.2	Limitations	163
8.3	Further study	164
References		165
Appendix A: Publications during the Ph.D study.....		185

Table of Figures

Figure 2-1 A flow chat of information retrieval/filtering.....	16
Figure 2-2 ‘Personalized recommendation’ page in Amazon	20
Figure 2-3 Link personalization from Excite’s Mylink page	21
Figure 2-4 Implicit content personalization in Amazon website.....	22
Figure 2-5 Structure personalization in My Yahoo! page.....	23
Figure 2-6 Interface of InfoFinder	34
Figure 2-7 An example of the movieLens system interfaces	41
Figure 2-8 An example of system interfaces in the Entrée Chicago	42
Figure 2-9 An example of personalized recommendations in Amazon.....	44
Figure 2-10 System interface of Electronic Funding Information.....	45
Figure 2-11 An example of system interface of Expertise Recommender	48
Figure 3-1 Research steps of the study	71
Figure 3-2 Methods and approaches used in this study	82
Figure 4-1 DeLone and McLean 's model for information system evaluation.....	85
Figure 4-2 An evaluation model of government online services.....	86
Figure 4-3 A development model of personalized government online services	99
Figure 5-1 Types of CF algorithms.....	105
Figure 5-2 The process of user-based similarity computation	108
Figure 5-3 The process of item-to-item similarity computation.....	109
Figure 5-4 Prediction process of item-based CF algorithm	112
Figure 5-5 Category of recommendation techniques in this study	113
Figure 6-1 A hierarchical product taxonomy tree.....	129
Figure 6-2 An example of the SPR model	131
Figure 6-3 A conceptual framework of dynamic personalization	133
Figure 6-4 Flow chart of recommendation process	136
Figure 6-5 Recommendation procedure.....	137
Figure 6-6 The architecture of STEF recommender system	142
Figure 6-7 Framework of SSA-based search engine	143
Figure 6-8 Hybrid fuzzy recommendation framework	145
Figure 6-9 An example of the STEF system interface.....	146
Figure 6-10 Member login page	146

Figure 6-11 User rating page.....	147
Figure 7-1 Impact of neighbor number on recommendation accuracy MAE	153
Figure 7-2 Impact of neighbor number on recommendation quality F1	153
Figure 7-3 Impact of β on recommendation accuracy MAE.....	154
Figure 7-4 Impact of β on recommendation quality F1	154
Figure 7-5 Comparison of maximum MAE values	155
Figure 7-6 Improvement extent of sparsity problem	156
Figure 7-7 Improvement extent of new item problem	157

List of Tables

Table 2-1 A list of types and techniques of main search engines of the web.....	15
Table 2-2 Examples of relevant techniques for improving CF approach	37
Table 2-3 Taxonomy of hybrid types.....	43
Table 4-1 List of constructs, measures and labels	87
Table 4-2 Comparison of Internal US and External US	92
Table 4-3 Reliability analysis with Cronbach's α coefficient.....	92
Table 4-4 ANOVA results to test relationship between constructs and internal US	93
Table 4-5 ANOVA results to test relationship between factors and external US	94
Table 5-1 User-item rating matrix	117
Table 5-2 Exhibition similarity degree matrix	118
Table 5-3 Exhibition semantic similarity matrix	118
Table 5-4 Total exhibition similarity matrix	118
Table 5-5 Prediction values of each exhibition	119
Table 6-1 Industry class from Austrade	130
Table 6-2 Example of products in the different levels.....	130
Table 6-3 The symbols used in data flow diagrams, their meaning and examples	135

Abstract

With the integration of information from different government agencies, a vast resource of information and services may be gathered in one portal. Many businesses have difficulty locating the required information and services. In such a situation of vast information overload, one of the difficulties facing governments is how to provide businesses with information specific to their needs, rather than an undifferentiated mass of information. One way to do this is through the development of personalized government online services. Indeed, the recent Accenture e-government study indicates that personalization techniques in e-government are beginning to emerge. However, existing personalization with recommendation techniques focuses on text document retrieval and e-commerce product recommendation domain. Personalization and recommendation applications in e-government have paid relatively little research attention.

Many mechanisms have been developed to deliver only relevant information to web users and prevent information overload. The most popular recent developments in the e-commerce domain are the user-preference based personalization and recommendation techniques. The existing techniques have a major drawback: they are difficulty to generate recommendation on one-and-only items, because most of them do not understand the item's semantic features and attributes. Therefore, this study aims to: (1) propose a novel approach, semantic product relevance model and its attendant personalized recommendation technique, to handle the one-and-only item recommendation problem; (2) develop a recommender system prototype, called Smart Trade Exhibition Finder, to tailor the relevant trade exhibition information to each particular business user, and to assist export business selecting the right trade exhibitions for market promotion. Smart Trade Exhibition Finder may reduce significantly the time, cost and risk faced by exporters in selecting, entering and developing international markets. In particular, the proposed approach can be used to overcome the drawback of existing recommendation techniques and enable recommender systems to work within a much wider range of problems which cannot currently be handled. The outcome of this study will solve the rating data lacking and new item problem, and significantly improve the performance compared to existing recommendation techniques.